

Building a bridge between Classical and Quantum Mechanics

Peter Renkel

January 18, 2017

Abstract

The way Quantum Mechanics (QM) is introduced to people used to Classical Mechanics (CM) is by a complete change of the general methodology (e.g. [1]) despite QM historically stemming from CM as a means to explain experimental results. Therefore, it is desirable to build a bridge from CM to QM.

This paper presents a generalization of CM to QM. It starts from the generalization of a point-like object and naturally arrives at the quantum state vector of quantum systems in the complex valued Hilbert space, its time evolution and quantum representation of a measurement apparatus of any size. It is shown that a measurement apparatus is a special case of a general quantum object.

1 Introduction

The basic problem with how QM is introduced is the fact that new definitions and relations between them are presented without keeping strong connection with CM. Indeed, QM is a more general theory than CM, and the latter can be obtained in the limit of the former. However, we live and interact with the classical world and our measurement apparatuses are mostly classical. Therefore, it is important to understand how we build QM starting from CM and generalizing it. Generalization is often not unique so the generalization we choose should be shown and justified.

Another problem related to the above is how measurement is introduced. In spite of the fact that it is central to QM, usually one loosely defines an apparatus as a macroscopic quantum object, and without clarifying what one means by 'macro' and why it is important, one jumps into Quantum Postulate [2]. This postulate contradicts the unitarity of the Evolution Operator, and hence all kind of paradoxes like 'Schrödinger's cat paradox' arise and give rise to multiple interpretations.

In this paper, I try to build the bridge between CM and QM. I obtain each QM postulate or a definition as a generalization of existing ideas in CM. Then I obtain QM parameters by matching CM to QM. For objects in QM that don't have corresponding matches in CM, I show the way of how to arrive at them. I use Occam's razor [4] to choose the simplest generalization if several generalizations are possible. If I don't know how to avoid arbitrariness in a specific case, I try to explicitly pose a question. I give a special attention to

the Problem of Measurement as central to QM. I try to avoid arbitrariness of the Quantum Postulate by explaining what is an apparatus, how to arrive at its definition and why it plays a special role in a bridge between QM and CM. I also try to avoid interpretations and present only the facts that follow from generalization of the CM.

2 Basics

The discussion below equally well applies to point-like objects and systems, characterized by generalized coordinates and momenta, or generalized coordinates and first derivatives of the coordinates. But starting from coordinate representation will give us an opportunity to easier map the actual dynamics to reality.

1. In CM we start with a notion of a point-like object. In CM, it is characterized by two vectors $\{\vec{x}, \vec{p}\}$. There is seemingly some arbitrariness in the number of vectors characterizing the point-like object. Why do we need to use two vectors and not three or one?
2. In QM we generalize this notion by saying that a point-like object is smeared over the whole space. Since we greatly increased the number of degrees of freedom for the object, let us see if we still need momentum as an additional degree of freedom¹. Thus, we start by providing the "density" of such an object in point \vec{x} as $\rho(\vec{x})$ ². The integral of such a function over the whole space is 1 corresponding to 1 object:

$$\int \rho(\vec{x}) d^3x = 1 \quad (1)$$

In CM, such a density is always a delta function of coordinates centered around the point where this object is located³.

3. In CM, since \vec{x} is a vector in a R^3 space, its coordinates will change depending on the system of coordinates, the orts are given by the axes of the system of coordinates.
4. In QM, our density $\rho(\vec{x})$ is a real-valued function. This function can also be represented as a vector in the real Hilbert space, where the basis is given by a particular set of orthonormal functions (e.g. delta functions). The coordinates (or coefficients of expansion) will change depending on the choice of the basis functions. Considering vectors, we have to define an invariant - a number that doesn't change when changing system of coordinates. The obvious candidate is vector's norm or the sum of squares of the coefficients. What is the physical quantity for an object that should be invariant? The simplest candidate is the number of objects. Therefore, we state that our candidate vector for one object is normalized and has norm 1 corresponding to one object.

$$d(\vec{x}) = \sum C_i d_i(\vec{x}), \quad (2)$$

¹Therefore, we won't utilize the Wigner distribution [3]

²It has nothing to do with the density matrix used in quantum statistics.

³As we will show, the density does not provide all information, but further generalizations will yield a consistent theory

where $d_i(\vec{x})$ are basis functions, C_i are the coefficients of expansion,

$$\sum_i (C_i)^2 = 1 \quad (3)$$

and $d(\vec{x})$ is a distribution function. For $d_i(\vec{x}) = \delta(\vec{x} - \vec{x}_i^0)$

$$\int d^2(\vec{x}) d^3x = 1 \quad (4)$$

From 1 and 4, it follows that

$$d(\vec{x}) = \pm \sqrt{\rho(\vec{x})} \quad (5)$$

So to make the number of objects invariant, we should move from density $\rho(\vec{x})$ to a new function - a distribution function defined by equation 5. Since density is a vector in the Hilbert space, we can completely get rid of index \vec{x} . Therefore, we are now associating a state vector $|D\rangle$ in real Hilbert space with a particle.

For an arbitrary expansion:

$$|D\rangle = C_1 |D_1\rangle + C_2 |D_2\rangle, \quad (6)$$

where $|D_1\rangle$ and $|D_2\rangle$ are two normal vectors, not necessarily orthogonal to each other:

$$1 = \langle D|D\rangle = C_1^2 + C_2^2 + 2C_1C_2 \langle D_1|D_2\rangle, \quad (7)$$

where $\langle D_1|D_2\rangle = \int d_1(\vec{x})d_2(\vec{x})d^3x$ is a scalar product of the two vectors.

3 A particle (system) with many independent degrees of freedom

1. In CM, a particle or a system with many independent degrees of freedom is characterized by a set of numbers. Each number corresponds to one degree of freedom.
2. In QM, by analogy we claim that a particle or a system is smeared over a $R^M \times H$ dimensional space, where M is the number of degrees of freedom. Thus, we define a vector for a particle in a specific point of this $R^M \times H$ dimensional space and get a $\rho(a_1, a_2, \dots, a_M)$, where a_i are independent characteristic parameters of a particle.

4 Multiple particles

1. Suppose we have N **non-interacting** particles. In CM, the generalization is straightforward: we have N vectors in R^{6^4} or 1 vector in $R^{6 \times N}$ space.

⁴Because in CM, we need to provide both coordinates and velocities, while in QM, it is enough to provide the state vector $\vec{x}_1 \dots \vec{x}_N$

- Generalization for QM is also straightforward. If we are given N non-interacting particles, we can consider each of them independently of all others. The state vector of the system is just a collection of respective state vectors⁵. To describe the system, where i^{th} particle has d_i degrees of freedom, we need $\sum_N d_i$ degrees of freedom.

If the order of the vectors in the collection matters, we can consider superpositions of collections with different order.

- What if an observer is not able to distinguish the particles? The density of two-particle state is constrained by: $\rho(\vec{x}_1, \vec{x}_2) = \rho(\vec{x}_2, \vec{x}_1)$. This in turn imposes the constraint on the distribution: $d(\vec{x}_1, \vec{x}_2) = \pm d(\vec{x}_2, \vec{x}_1)$. Then the density of the two-particle state can be written as:

$$d(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}}(d(\vec{x}_1)d(\vec{x}_2) \pm d(\vec{x}_2)d(\vec{x}_1)) \quad (8)$$

Since the particles don't interact, the scalar product of two distributions vanishes and the sign between two terms in equation 8 can be plus or minus. Several observations can be made here:

- The density does not depend on the sign in equation 8. Both signs yield the same expression for the density.
- In CM, the wave functions are effectively delta functions. Equation 8 then reads:

$$\rho(\vec{x}_1, \vec{x}_2) = \frac{1}{2}(\delta(\vec{x}_1 - \vec{x}_1^0) + \delta(\vec{x}_2 - \vec{x}_2^0)) \quad (9)$$

However, in QM equation 8 can lead to interference of two terms when squared but in CM it won't because delta functions are orthogonal to each other.

- In future, an observer might find a way to distinguish particles. Then, he should use the full distribution. By examining if equation 8 works, an observer can test if the particles under investigation are identical. It would be interesting e.g. to derive identity of protons starting from quarks.
 - The source of equation 8 is **invariance of the observable density**, not distribution, since the latter is unphysical.
2. Suppose we have N **interacting** particles. In CM, the system is described by N vectors, time evolution of each of them depends on the positions and momenta of the others. We can also describe the system with 1 vector in $R^{6 \times N}$ space corresponding to N particles but the evolution of this vector is complicated.
- If we try to naively generalize this understanding to QM, we run into a problem. Since the particles are now interacting, the fact that we know the density distribution for one of them gives us some information about the others. Say for an attractive force, we expect the densities of the particles be peaked roughly in the same area. Hence there is a correlation between the densities of particles.

⁵Generally speaking, the set of all such collections doesn't constitute a linear space

Let's consider a system of two interacting particles. Suppose, they both have distributions:

$$d(x) = C_1 \exp\left(-\frac{(x - x_0^1)^2}{2\sigma^2}\right) + C_2 \exp\left(-\frac{(x - x_0^2)^2}{2\sigma^2}\right), \quad (10)$$

where C_1 and C_2 are some normalization coefficients, x_0^1 and x_0^2 correspond to peaks and σ correspond to widths of the two Gaussians⁶. But we expect the distribution of the second particle to peak at the same places where the distribution of the first particle peaks (points x_0^1 and x_0^2), otherwise we will lose the information about correlations. Therefore, it is not enough to know the distributions of each particle, we also need to provide their correlations.

The solution is to introduce the density of the pair:

$$\rho(x, y) = D_1 \exp\left(-\frac{(x - x_1^1)^2}{2\sigma^2}\right) \exp\left(-\frac{(y - y_1^1)^2}{2\sigma^2}\right) + D_2 \exp\left(-\frac{(x - x_1^2)^2}{2\sigma^2}\right) \exp\left(-\frac{(y - y_2^2)^2}{2\sigma^2}\right) \quad (11)$$

Notice, writing the density of the pair in such form, **we gain the knowledge about correlations, but we loose identities of the particles** since it is impossible now to identify the distribution of a specific particle in the mixture. In fact, a distribution for a single particle can exist only if it doesn't interact with anything else and didn't interact before, because if it did, correlations became part of the multi-particle density. Generally, the simplest way to account for correlations is to write the multiparticle state vector as:

$$|D\rangle = \sum_{i_1, i_2, i_3, \dots, i_N} C_{i_1, i_2, \dots, i_N} |D_{i_1}\rangle |D_{i_2}\rangle \dots |D_{i_N}\rangle, \quad (12)$$

where all summation indices are different. Now vectors $|D\rangle$ form a linear space - a direct product of vectors spaces corresponding to each particle. Please note, correlation coefficients C_{i_1, i_2, \dots, i_N} can change in time when a system evolves.

This situation appears only in QM since in CM we precisely know the positions and momenta of the particles and hence we should not worry about correlations.

- For the identical interacting particles, the sign in equation 8 is now meaningful since the density gets another term that depends on the cross-product of the two particles.

$$\rho(\vec{x}_1, \vec{x}_2) = \rho(\vec{x}_1)\rho(\vec{x}_2) \pm 2| \langle d(\vec{x}_1) | d(\vec{x}_2) \rangle |^2 \quad (13)$$

However, swapping the two particles yield the same density. For N particles, we require that the N particle density doesn't change under any permutation and arrive at the Slater determinant [10].

⁶For simplicity, we assume the Gaussians are symmetric in x-y plane.

- The notion of interacting particles as it is stated above does not have any practical value because in order to trace the evolution of a system, we need to consider the time evolution of the whole Universe - no weak interactions can be ignored. For the practical purposes, we need to be able to perform some sort of perturbation theory for weak interactions and maintain densities for our system. There are cases when we can do it by introducing potentials acting on a system. However, we can separate the distributions of the system and the rest of the Universe only under specific conditions. To be able to find these conditions, we need to investigate how a closed system evolves in time.

5 Time Evolution

- In CM, the coordinates of a particle can be described by $\vec{x}(t)$ - a **function** of t . Also, there is a function that takes coordinates at time t_0 and gives us new coordinates at time t .
- In QM, the $d(\vec{x}_0, t)$ is a function of t for a given \vec{x}_0 . Consequently, there is an operator acting on initial function (or a vector in Hilbert space) at t_0 to give us a function (or a vector in Hilbert space) at time t . It is not easy to generalize CM Newtonian laws for such a case. So let us start by stating the minimal set of meaningful requirements for such an operator. We will use Occam's razor, so if we could impose a range of constraints, we use the simplest. After stating such requirements and finding an operator satisfying them, we should make sure that we can get CM in the limit when above mentioned densities are delta functions.
 1. This operator preserves the norm (we stated that our $d(\vec{x})$ is always normalized).
 2. This operator should be linear. This is the simplest non-trivial property of operators we may impose.
- The simplest non-trivial linear transformation that preserves the norm is just a rotation - an orthogonal transformation. To derive the exact form of this operator, we first need to investigate how two quantum systems interact with each other.

6 Measurement

Suppose we perform a measurement of a system - an interaction of our test system with a probe system. This means that we take two non-interacting sub-systems (test and probe) and start turning on interactions.

- In CM, initially the composite system is described by two vectors - one for the probe and one for the test system. After interaction, the probe doesn't change the test system - it just measures it's state.
- In QM, initially, the composite system is described by the collection of two vectors (corresponding to two systems), but eventually the vector evolves into a vector in the direct product space.

Since in CM, test systems don't change after interaction, let's try to keep this constraint and consider the set of vectors of the test systems that are multiplied by a real number after the interaction⁷ [6]. We can show that these vectors are orthogonal. For two such vectors u_i and u_j where $i \neq j$:

$$\langle u_i | u_j \rangle = \langle u'_i | u'_j \rangle = \alpha_i \alpha_j \langle u_i | u_j \rangle, \quad (14)$$

where u'_i and u'_j are the vectors after interaction, α_i and α_j are some real numbers. Here we used the fact that the scalar product is preserved under orthogonal transformations.

$$\langle u_i | u_j \rangle (1 - \alpha_i \alpha_j) = 0. \quad (15)$$

Since α_i and α_j are arbitrary and there is no reason to expect their inner product to be 1,

$$\langle u_i | u_j \rangle = 0 \quad (16)$$

We can also define a linear orthogonal operator (\hat{O}) for which these vectors are eigenvectors. We call this operator - a measurement operator. This is the same operator that is responsible for evolution of individual vectors.

Vectors u_i are delta functions (or delta symbols) in the eigenbasis of the measurement operator ($\delta(x - x_i^0)$, where x_i^0 are the corresponding eigenvalues). This means that their eigenvalues correspond to the centers of the delta functions, and that's why they provide the **measured** values of the corresponding variables. This fact is stated in QM books but only now its physical sense becomes clear.

Generally speaking, eigenvectors of \bar{O} are complex valued. In section 2 we started by requiring our $d(\vec{x})$ to be real, however now we see that:

- We either need to restrict the set of our measurement operators and coordinates, so that the eigenvectors are always real.
- Or we need to abandon the requirement that $d(\vec{x})$ is real.

It is not easy to satisfy the first requirement but the second requirement, although seems problematic from the physics perspective, actually is not. Indeed, as we found above, the only thing we measure are eigenvalues, and these are real. So we can start by not imposing any extra requirements and allowing the $d(\vec{x})$ to be complex. We will use $\psi(\vec{x})$ for what we called distribution function before and call it 'the wave function'. Also, since our wave function is complex, there is no sense of keeping the measurement operator orthogonal. We will use unitary operators for the measurement operators instead of orthogonal operators. Now, a particle is described by a vector (we also call it a state vector) in a complex Hilbert space. The norm of this vector is 1. The most obvious choice for the above density is $|\psi(\vec{x})|^2$. All earlier discussions concerning orthogonal operators stay intact since orthogonal operators are a subset of unitary operators.

For a wave function of N identical non-interacting particles, adding an arbitrary phase in front of any term will not change the density. However, for N identical interacting particles, we should make sure that any permutation does

⁷The corresponding normalized vector characterizing the state of the system does not change.

not change the density. It is easy to see that this can be accomplished only for the Slater determinant [10].

Here is the summary of what we achieved so far.

1. We started by introducing a function $\rho(\vec{x})$ corresponding to the density of a particle, such that $\int \rho(\vec{x}) d^3x = 1$.
2. This allowed us to represent a particle as a vector in Hilbert space. We then concluded that in order for the number of particles to be invariant, we need to change our definition to that in equation 5.
3. In this section we realized that the $d(\vec{x})$ can also be complex value and called it $\psi(\vec{x})$. Now $\psi(\vec{x})$ does not have a direct physical meaning, but $|\psi(\vec{x})|^2$ corresponds to the density we started from.

7 Time evolution operator

Let's get back to the evolution operator and see if we can figure out its exact form.

- Since now we have a complex-valued function, our rotations will be represented by unitary transformations : $\sum_j \exp(iA_{kj})\psi_j$, where A_{kj} is a hermitian matrix.
- Let us diagonalize this matrix. We will get $A_{kj} \rightarrow D_{kk} * \lambda_k$, where λ_k are the eigenvalues of the operator A . They can be interpreted as angles of rotations around corresponding axes k . The simplest non-trivial dependence of such rotation on time is rotation with constant angular velocities. Therefore we state that $\lambda_k = \omega_k * t$, where ω_k is a constant. The evolution in such a basis will be given by:

$$\psi'_k(t) = \psi'_k(0) \exp(-i\omega_k t), \quad (17)$$

where $\psi'_k(t)$ is the k 's component of $\psi'(t)$ in the system of coordinates where A_{jk} is diagonal. It is clear from above equation that the state vector 'turns' the k 's coefficient c_k with constant angular velocity ω_k .

- The evolution of the state vector then takes the form :

$$\psi(t) = \psi(0) \exp(-i\hat{A}t) \quad (18)$$

For $-\hat{A}\hbar$ being Hamiltonian, we get the Schrödinger equation.

- Let's expand the state vector. Previously, we were talking about a specific set of eigenfunctions - delta functions:

$$\psi(\vec{x}, \vec{x}_i^0) = \psi(\vec{x})_{\vec{x}_i^0} = \delta(\vec{x} - \vec{x}_i^0), \quad (19)$$

but there is infinitely many possible bases, each yields a description of our system. Another basis is a set of harmonic functions:

$$\psi_k(x) = C \exp(i\vec{k}\vec{x} + \phi), \quad (20)$$

where C is a normalization constant. The phase ϕ is just an overall constant factor and can be absorbed into coefficient C .

Let's introduce a new coordinate p and define it:

$$\vec{p} = \hbar \vec{k} \quad (21)$$

where \hbar is an arbitrary number called "Planck's constant". This transformation is just a scale to use more convenient units for measuring \vec{k} . Transition from coordinate representation to a new representation is then given by:

$$\psi(\vec{p}) \sim \int \exp\left(i \frac{\vec{p} \cdot \vec{x}'}{\hbar}\right) \psi(\vec{x}') d^3 x' \quad (22)$$

If we define an operator \hat{p} such that

$$\hat{p} = -i\hbar \nabla, \quad (23)$$

we get using equation 22:

$$\psi(\vec{p}) = \hat{p} \psi(\vec{x}) \quad (24)$$

At this point, we suspect that \vec{k} is related to one of our fundamental variables in QM. The closest analogy in CM is wave vector or momentum⁸. By matching QM dynamics to CM, we will be able to see in section 11 that this is indeed the case.

8 Maximum entanglement and decoherence

Let's come back to interactions of the two subsystems - test and probe:

$$|\psi\rangle = \sum_i C_i |u_i\rangle |\chi\rangle \quad (25)$$

The test subsystem is expanded to the basis of orthonormal vectors above, and χ is the wave function of the probe subsystem before the measurement. After the measurement we have:

$$|\psi'\rangle = \sum_i D_i |u_i\rangle |v_i\rangle, \quad (26)$$

where D_i are complex value coefficients, u_i are the test eigenfunctions and v_i are some vectors in the second subsystem. Note that u_i didn't change. For an arbitrary interaction we can assume that v_i are random normalized vectors in Hilbert space. Let's calculate $\langle v_i | v_j \rangle$ where $i \neq j$. For two arbitrary vectors in a N dimensional space, we can think of their inner product as a projection of v_j on v_i . For a random vector v_j , we assume that all projections are equal on average. Thus, $v_1 = v_2 = \dots = v_n$ and hence since $\sum_i v_i^2 = 1$, on average:

$$\langle v_i | v_j \rangle = 1/\sqrt{N}. \quad (27)$$

⁸According to equation 22, \vec{k} has dimension $1/m$. If we want it to be momentum, our multiplicative parameter should have dimension $\frac{1/m}{m/s} = \frac{s}{m^2}$.

For $N \gg 1$,

$$\langle v_i | v_j \rangle \sim 0. \quad (28)$$

For the second subsystem with many degrees of freedom, we conclude that the inner product is zero. When v_j are almost orthogonal, we say that the system decohered, and we call the whole process - decoherence [6]. Equation 26 with orthonormal u_i and v_j are called Shmidt decomposition [5]. The time it takes a system to decohere depends on the number of degrees of freedom and mass of the subsystems, and generally is very small (below what we can measure) for large systems (see equation 27). However, for say an interaction of an electron with another electron, this time is long.

For large probes, $|v_i\rangle$ in equation 26 are orthogonal. Since both $|u_i\rangle$ and $|v_i\rangle$ form an orthogonal set, each term $D_i |u_i\rangle |v_i\rangle$ is completely independent of all others. Therefore, after the interaction, the system consists of N completely independent terms - we will call them 'branches'.

One immediate consequence is that once two subsystems interacted, they are always mixed, we won't be able to find individual subsystems in the mixture. It might mean that there are no unmixed systems in the world and everything is entangled with everything. The wave function of the Universe then can be represented by:

$$\psi_{universe} = \sum_{branches_i} \prod_{objects_j} |v_{branch_i, object_j}\rangle, \quad (29)$$

where $\langle \prod_{objects_j} v_{branch_{i1}, object_j} | \prod_{objects_{j'}} v_{branch_{i2}, object_{j'}} \rangle = 0$ for objects that decohered.

9 Measurement and the Quantum Postulate

We notice an important consequence of the above discussion : unlike in CM where we can always outline subsystems in the composite system, in QM the subsystems are mixed. Moreover, according to Shmidt decomposition, each system having interacted with another large system gets entangled. From this point in time, the state vectors of individual systems don't exist anymore even when the systems are far from each other and not interacting.

Let's define an apparatus as a quantum system which decoheres fast. Usually apparatuses are large and have many degrees of freedom. After a particle interacts with an apparatus, the wave function of the apparatus-particle composite system is given by equations 26 and 28. There are multiple completely independent branches in this decomposition.

Let's take a concrete example : a double slit experiment followed by two films parallel and very close to each other⁹. Let's assume that an electron penetrates the first film leaving a spot on the film, and then gets absorbed by the second film. There is also an observer (who has many degrees of freedom), observing the spots on two films.

⁹In order for the electron wave functions not to evolve considerably while electron travels between the films

The state vector of the system after an electron gets absorbed by the second film looks like.

$$\langle \vec{x} | system \rangle = \int \langle \vec{x} | (|ele\rangle |film_1\rangle |film_2\rangle |\gamma\rangle |obs\rangle) d^3x, \quad (30)$$

where $|\gamma\rangle$ is a photon state vector that delivers information to observer's eyes¹⁰, and \vec{x} is an index corresponding to a coordinate along film's axes. Also, $\langle x|ele\rangle$, $\langle x|film_1\rangle$, $\langle x|film_2\rangle$ are delta functions centered at coordinate \vec{x} . Please note, the observer is no longer decoupled from the apparatus, moreover, it is entangled with the $e - film_1 - film_2 - \gamma$ system. Therefore, it is wrong to imagine an observer looking at the system because the observer and the system don't exist anymore, but the entangled system does.

We should read equation 30 as continuum of observers $obs_{\vec{x}}$ each looking at the corresponding $\langle x|ele\rangle$, $\langle x|film_1\rangle$, $\langle x|film_2\rangle$. Each observer with index \vec{x} will see a spot on the first film at point \vec{x} and on the second film at the same point \vec{x} . Also if an observer \vec{x} sees a spot at \vec{x} and sees another observer, then another observer also sees the spot at the same point \vec{x} . In other words, there is no collapse, wave function evolves according to Schrödinger equation, however the appearance of collapse is achieved due to entanglement of the observer with the system.

Let's take another example : EPR experiment [7]. In this experiment, the pair of particles are not only entangled with each other but are also entangled with the observer. Therefore, one pair of observers would see the first branch, and another pair would see another branch. No superluminal velocities are involved since the branches existed even before the measurement.

10 Born's rule

In section 9 we realized that a particular observer can't see the whole wave function of a subsystem simply because this wave function doesn't exist for the subsystem, instead the wave function of the whole system exists. However, an observer attached to a particular branch, can see a system corresponding to this branch.

Now the natural question arises: "What is the probability that a particular observer finds himself in a branch with number i observing a subsystem i ?" This question cannot be addressed by Schrödinger equation alone since this equation gives the time evolution of the wave function of the whole system, and doesn't provide any information about probabilities¹¹. At this point, we need to return to our original notion of density. We introduced $\rho(\vec{x})$ as a density of our electron smeared all over the space. Since then, there were two significant changes:

1. We introduced a complex valued wave function
2. This wave function exists only for non-interacting systems which were never entangled, and it is wrong to assign wave function to a subsystem after interaction.

¹⁰Since photons are relativistic objects and their number does not conserve, the wave functions of photons are strictly speaking not defined

¹¹Indeed, so far we found out that wave function is related to the density but never discussed its physical meaning. Schrödinger equation is just an equation that takes the wave function at time zero, and returns the wave function at time t

The second change forces us to seek for the probabilities corresponding to a wave function of the system which was never entangled. But this function is complexed values.

Let's consider M sub-systems after they all decohered. Suppose we have N branches in the wave function of the composite system::

$$|\psi_{system}\rangle = \sum_{i=1...N} \rho_i \exp(i\phi_i) |\psi_i^1\rangle > |\psi_i^2\rangle > ... |\psi_i^M\rangle, \quad (31)$$

where ρ_i is the absolute value of the i complex coefficient and ϕ_i is its phase. Also, ϕ_i is the sum of all phases for all sub-systems.

Let's recall how we started. Our $\psi(\vec{x})$ came from the distribution $d(\vec{x})$ whose square defines a density in a point \vec{x} . We can just state that our system is smeared over the branches - each branch gets a complex weight equal to the corresponding density. The density for the branch is then: $\rho_i(\vec{x}) = |\psi(\vec{x})|^2$. Since each object has observers attached to it, we can expect the density to be proportional to the number of observers attached to this branch or the probability for an observer to find himself attached to the corresponding branch. This constitutes the Born's postulate.

We also need to prove that our definition of the probabilities does not create contradictions down the line. In [6], a derivation of Born's postulate from symmetry consideration is provided which can also serve as such a prove.

11 Mapping to CM and exact form of the evolution operator.

In CM, we use equations of motion to trace the evolution of the system. In QM, as we found out in section 5, we use the time evolution operator. It is important to obtain CM starting from QM and taking some limit. We will define the limit as follows: suppose we measure some property of a system multiple times¹². Each time after a measurement, we get some measured value. Averaging these values will give us some value. For a classical system, this value doesn't change and equals the measured value.¹³ Therefore, for classical systems, QM averages and equations of motion coincide with classical values and equations of motion.

According to the previous section, each time measuring the system we just get its state vector corresponding to one of the branches. When we average the results, we give a weight to each branch, corresponding to the Born's rule [6]. Thus, we arrive at the well known form for expected values which maps QM to CM.

$$M = \langle \psi | \hat{M} | \psi \rangle, \quad (32)$$

where \hat{M} in the *rhs* of equation 32 is an operator of the apparatus that measures M - its classical value.

For interacting subsystems, generally speaking we cannot separate each subsystem from the many-particle wave function. Therefore, when trying to match

¹²Of course, after the measurement, we create a mixed state of the system with the apparatus so it is practically impossible to get back to the initial state vector. So in practice we prepare many systems with identical state vector and let them interact with the set of identical apparatuses.

¹³Actually, there will be some scatter due to the imperfect apparatus, but this scatter is not related to the quantum mechanical scatter.

QM to CM, we should start from systems not entangled with anything and consider their evolution. Suppose we have two interacting particles, the classic potential between them is given by $U(x_1, x_2)$.

We know that the classical energy for a closed system, total momentum and angular momentum conserve:

$$\frac{dE}{dt} = 0 \quad (33)$$

$$\frac{d\vec{p}}{dt} = 0 \quad (34)$$

$$\frac{d\vec{L}}{dt} = 0 \quad (35)$$

Let us use equation 18 to find the evolution of the expectation value for the variable $M(t)$:

$$\begin{aligned} \frac{dM(t)}{dt} &= \frac{d \langle \psi(x_1, x_2, t) | M \rangle | \psi(x_1, x_2, t) \rangle}{dt} = \\ &= \frac{d \left\langle \psi(x_1, x_2, 0) \exp\left(-\frac{i\hat{A}t}{\hbar}\right) \middle| M \right\rangle \left| \exp\left(\frac{i\hat{A}t}{\hbar}\right) \psi(0) \right\rangle}{dt} = \\ &= \left\langle -\frac{i\hat{A}}{\hbar} \psi(0) \exp\left(-\frac{i\hat{A}t}{\hbar}\right) \middle| M \right\rangle \left| \exp\left(\frac{i\hat{A}t}{\hbar}\right) \psi(x_1, x_2, 0) \right\rangle + \\ &= \left\langle \psi(x_1, x_2, 0) \exp\left(-\frac{i\hat{A}t}{\hbar}\right) \middle| M \right\rangle \left| \exp\left(\frac{i\hat{A}t}{\hbar}\right) \psi(x_1, x_2, 0) \frac{i\hat{A}}{\hbar} \right\rangle = \\ &= \frac{i}{\hbar} \langle \psi(x_1, x_2, 0) | [\hat{M}\hat{A}] | \psi(x_1, x_2, 0) \rangle \quad (36) \end{aligned}$$

This means that if we choose:

$$\hat{A} = f(\hat{H}), \quad (37)$$

or

$$\hat{A} = f(\vec{\hbar}_p)\hat{p}, \quad (38)$$

or

$$\hat{A} = f(\vec{\hbar}_L)\hat{p}, \quad (39)$$

where \hat{H} is the operator corresponding to energy and $\vec{\hbar}_p, \vec{\hbar}_L$ are some constant vectors, we will restore conservation of energy or momentum in classical limit. Choosing $\hat{A} = \vec{\hbar}_p$ or $\hat{A} = \vec{\hbar}_L$ breaks rotational invariance so using Occam's razor, let's choose:

$$\hat{A} = \hat{H} = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m_2} + U(\hat{x}_1, \hat{x}_2) = -\frac{\hbar^2}{2m_1}\Delta - \frac{\hbar^2}{2m_2}\Delta + U(\hat{x}_1, \hat{x}_2), \quad (40)$$

and check if the total momentum and angular momentum also conserve. We can use our new form for the time evolution operator to recover the Ehrenfest

equations [9] for values of momentum p and coordinate x of the test particle¹⁴.

$$\frac{d\vec{p}_1}{dt} = -\nabla_1 U(\vec{x}_1, \vec{x}_2), \quad (41)$$

$$\frac{d\vec{p}_2}{dt} = -\nabla_2 U(\vec{x}_2, \vec{x}_1), \quad (42)$$

$$\frac{d\vec{x}_1}{dt} = \frac{\vec{p}_1}{m} \quad (43)$$

$$\frac{d\vec{x}_2}{dt} = \frac{\vec{p}_2}{m} \quad (44)$$

The conservation of total momentum $p_1 + p_2$ follows from CM.

It is worthwhile to notice that in CM, Planck's constant disappears, it cancels out. However, an exact value of \hbar in QM actually defines our time evolution in 40 and 23.

In many textbooks, Schrödinger equation is postulated making a reader wonder why Schrödinger equation takes this and not any other form. We tried to arrive at the Schrödinger equation naturally and prove its consistency with CM.

12 Weakly interacting systems

In Section 4, we were trying to find such weakly interacting subsystems of the bigger subsystems, for which each of the subsystems could be identified as a separate subsystem subject to Schrödinger equation with some potential.

However, equation 29 tells us that everything is entangled with everything, and there is no way to factor our single particles from the wave function of the Universe. How to resolve this contradiction?

Equation 30 implies that the wave function consists of orthogonal branches, and each branch has a set of observers attached to it. These observers see only this branch and for them, the particles are not entangled - the individual particles can be identified.

After some time however, if the particles interact, the observers will not be able to find individual particles in the corresponding branch. So the above question about weakly interacting subsystems should be posed for one branch. Initially the branch consists of individual particles. Under which conditions after time t , the particles can still be identified? If these conditions are not satisfied, the branch will again branch out to sub-branches - each sub-branch will have a new set of observers attached to it.

There are several important conditions for preventing sub-branching:

1. Let's consider two interacting particles. Their interaction is described by a potential $U(\vec{x}_1, \vec{x}_2)$, where \vec{x}_1 is the coordinate of the first particle, and \vec{x}_2 is the coordinate of the second particle. The Schrödinger equation can be written as follows:

$$i\hbar \frac{\partial \psi(\vec{x}_1, \vec{x}_2, t)}{\partial t} = \left(\frac{\vec{p}_1^2}{2m_1} + \frac{\vec{p}_2^2}{2m_2} + U(\vec{x}_1, \vec{x}_2) \right) \psi(\vec{x}_1, \vec{x}_2, t) \quad (45)$$

¹⁴The individual wave functions don't exist, but we can still find the average properties for each particle by averaging the corresponding operators using two-particle wave function

For the linear potentials of the form $U(\vec{x}_1, \vec{x}_2) = \alpha|\vec{x}_1 - \vec{x}_2|$, the Schrödinger equation can be factorized:

$$i\hbar \frac{\partial \psi(\vec{x}_1, t)}{\partial t} = \left(\frac{\vec{p}_1^2}{2m_1} + \alpha\vec{x}_1 \right) \psi(\vec{x}_1, t) \quad (46)$$

$$i\hbar \frac{\partial \psi(\vec{x}_2, t)}{\partial t} = \left(\frac{\vec{p}_2^2}{2m_1} - \alpha\vec{x}_2 \right) \psi(\vec{x}_2, t), \quad (47)$$

where $\psi(\vec{x}_1, \vec{x}_2, t) = \psi(\vec{x}_1, t)\psi(\vec{x}_2, t)$. In this case (e.g. a particle in an electrostatic field), the particle retains its identity, but Schrödinger equation includes an external potential.

2. If a potential changes slowly, we can use only the first term in Taylor's expansion and approximately factor out multiparticle wave function. For a potential to change slowly, the second term in the Taylor's expansion should be much smaller than the first term. In other words, from Schrödinger equation:

$$\left| \frac{1}{2} \int \frac{\partial^2 U}{\partial r^2} \Big|_{r_0} (r - r_0)^2 \psi(r) dr \right| \ll \left| \int \frac{\partial U}{\partial r} \Big|_{r_0} (r - r_0) \psi(r) dr \right|, \quad (48)$$

where r is the coordinate from one particle to another. For a wave function with the width of σ , and potential of $\pm \frac{\alpha}{r}$, equation 48 reduces (up to coefficients) to:

$$\sigma \ll r \quad (49)$$

To be conservative, we can introduce a coefficient $\omega < 1$ representing the degree of linearity of our potential and write the above equation as:

$$\sigma = \omega r, \quad (50)$$

The important conclusion we can draw from equation 50 is that **we can talk about particles only when the widths of their individual wave functions are much smaller than the distance between them.**

3. For a particle in the constant magnetic field, only kinetic terms change, and the system can still be factorized.
4. Wave functions of classical objects are narrow. This means that positions and momenta of each object is set and there is no place for correlations, hence multiobject wave function can be factorized to yield a product of individual wave functions. Therefore, each classical object is a subject of classical equations of motions (see also Section 15).

However, generally an interacting particle cannot be separated from the environment. Say, for two particles wave function, we can change the variables : $\vec{x}_1, \vec{x}_2 \rightarrow \vec{r} = \vec{x}_1 - \vec{x}_2$. For the new variable, Schrödinger equations reads:

$$i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} = \left(\frac{\vec{p}^2}{2\mu} + U(\vec{r}) \right) \psi(\vec{r}, t), \quad (51)$$

where $\mu = \frac{m_1 m_2}{m_1 + m_2}$ is a reduced mass. However, despite the Schrödinger equation is now 1-dimensional, we didn't factorize the system, the particles are still entangled.

This very important fact is the source of the Quantum postulate and is overlooked in many QM discussions.

13 Composite systems

Suppose we have a system composed of many particles. How do we treat such a system? Should we consider equations of motion of each individual particle to figure out the behavior of the whole system?

- In CM, we prove that Newtonian equations of motion apply to the center of mass of the system, and we can ignore the fact that the system is composite when considering its evolution¹⁵.
- In QM, this question needs some clarification because as stated above, we can talk about standalone particles only under specific constraints. Suppose we have a system composed of many particles. This system is far enough from all other particles, so for each particle inside this system, equation 50. Can we treat this system as a particle, namely as having a wave function and evolving according to Schrödinger equation with some potential?

The answer to this question usually quietly assumed to be true, but it needs to be proved. For a system of N particles in a potential U , Schrödinger equation reads:

$$i\hbar \frac{\partial \psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t)}{\partial t} = \sum_{i=1..N} \left(\frac{\hat{p}_i^2}{2m_i} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \right) \psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t), \quad (52)$$

where

$$U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \sum_{i,j=1..N} V^{int}(\vec{r}_i - \vec{r}_j) + \sum_{i=1..N} V^{ext}(\vec{r}_i), \quad (53)$$

and V^{int} , V^{ext} are potentials corresponding to interactions between particles and interaction between the composite system and other particles respectively. Guided by CM solution, we want to factor out wave function of the center of mass of the system from the N -particle wave function:

$$\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = \psi(\vec{r}_C, t) \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t), \quad (54)$$

where:

$$\vec{r}_C = \frac{\sum_i m_i \vec{r}_i}{\sum m_i}. \quad (55)$$

Plugging equations 55, 54 and 53 into equation 52, we get for the wave

¹⁵For figuring out internal time evolution of the system (e.g. rotations), we need to take into account its internal interactions, but to get the evolution of the system as a whole, we can ignore them.

function of the center of mass:

$$\begin{aligned}
& i\hbar \frac{\partial \psi(\vec{r}_C, t)}{\partial t} \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) + \\
& i\hbar \frac{\partial \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t)}{\partial t} \psi(\vec{r}_C, t) = \\
& - \frac{\hbar^2}{2 \sum m_i} \Delta \psi(\vec{r}_C, t) \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) \\
& - \sum_i \frac{\hbar^2}{m_i} \nabla \psi(\vec{r}_C, t) \nabla \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) - \\
& \sum_i \frac{\hbar^2}{2m_i} \psi(\vec{r}_C, t) \Delta_i \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) + \\
& U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \psi(\vec{r}_C, t) \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) = \\
& - \frac{\hbar^2}{2 \sum m_i} \Delta \psi(\vec{r}_C, t) \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) - \\
& \sum_i \frac{\hbar^2}{2m_i} \psi(\vec{r}_C, t) \Delta_i \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) + \\
& U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \psi(\vec{r}_C, t) \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) \quad (56)
\end{aligned}$$

We can break this equation into two parts:

$$i\hbar \frac{\partial \psi(\vec{r}_C, t)}{\partial t} = - \frac{\hbar^2}{2 \sum m_i} \Delta \psi(\vec{r}_C, t) + \sum_i V^{ext}(\vec{r}_i) \psi(\vec{r}_C, t), \quad (57)$$

and

$$\begin{aligned}
& i\hbar \frac{\partial \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t)}{\partial t} = \\
& \sum_i \frac{\hbar^2}{2m_i} \Delta_i \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t) + \\
& \sum_{i,j=1..N} V^{int}(\vec{r}_i - \vec{r}_j) \psi(\vec{r}_1 - \vec{r}_C, \vec{r}_2 - \vec{r}_C, \dots, \vec{r}_N - \vec{r}_C, t). \quad (58)
\end{aligned}$$

It follows from equation 57 that when considering evolution of composite systems, we can decouple the wave function of its center of mass from the wave function connected to internal movements. Therefore, we can ignore particles inside these systems and just solve Schrödinger equation for their centers of mass. We need to point out that this is possible because the product of two gradients vanishes. If we were to choose a different Hamiltonian than 40, we might not been able to get this important property of the composite system, and in order to understand evolution of the composite objects, we would need to solve equations of motion for elementary objects. Clearly, our world is modular - generally speaking, we don't need to talk about electrons when considering stones.

14 Interaction of two particles

Let us consider two particles with masses m_{probe} and m_{test} running into each other. When the particles are far from each other so that equation 50 holds, they can be separated, and the evolution of each particles is described by Schrödinger equation. When the particles are close so that the width of their wave functions are larger than the distance between them, this approximation is invalid. Let us assume for simplicity that the wave function of the probe particle is much smaller than the distance between the particles. We can then expand the wave function of the test particle:

$$\psi^{probe} = \sum C_i^{probe} \psi_i^{probe} \quad (59)$$

where ψ_i^{probe} are chosen in such a way that for each ψ_i^{probe} the potential is approximately linear (equation 48). Thus, our Schrödinger equation for two-particle wave function can be approximated by a series of Schrödinger equations for two distinct particles.

$$i\hbar \frac{\partial \psi_i^{probe}(\vec{x}_t)}{\partial t} = \left(-\frac{\hbar^2}{2 \sum m_t} + V(\vec{x}_t - \vec{x}_p) \right) \psi_i^{probe} \quad (60)$$

and

$$i\hbar \frac{\partial \psi_i^{test}(\vec{x}_p)}{\partial t} = \left(-\frac{\hbar^2}{2 \sum m_p} + V(\vec{x}_p - \vec{x}_t) \right) \psi_i^{test}, \quad (61)$$

where $V(\vec{x})$ is a linear potential, and ψ_i^{test} is test wave function corresponding to the i^{th} wave function of the probe. After some time interval t_D different branches corresponding to equations 60 and 61 become almost orthogonal to each other. We can estimate this time interval using Ehrenfest equations 41-44 for the averages of the wave functions

$$\langle \psi_i^{probe}(0) \rangle \rightarrow \langle \psi_i^{probe}(t) \rangle = \langle \psi_i^{probe}(0) \rangle + \frac{F_i}{m_p} \frac{t^2}{2}, \quad (62)$$

where F is the classical force acting between the particles. After the decoherence time t_D the distance between two consecutive averages will equal the width of the probe functions, hence the probe wave functions become almost orthogonal to each other. For a force:

$$F = \pm \frac{q^2}{4\pi\epsilon_0 r^2}, \quad (63)$$

the decoherence time

$$t_D = \sqrt{\frac{2\sigma_p m_p}{F_i - F_{i-1}}} \quad (64)$$

is:

$$t_D = \sqrt{\frac{8\pi\epsilon_0\sigma_p m_t (r_i^2 - r_{i-1}^2)}{q^2}} = \sqrt{\frac{8\pi\epsilon_0\sigma_p m_t r_i^2 \omega^2}{q^2}} \quad (65)$$

For an electron hitting a proton separated by a distance of 1 Å assuming the width of the 1 Å for the proton, and taking ω as 0.01, we get $t_D \sim 10^{-17}s$. However, for the initial distance of $1m$, the time is $\sim 10^{-2}s$.

In reality, our testing and probe systems are coupled to the thermal bath which makes the task of calculating decoherence time harder. It is beyond the limits of this paper to calculate decoherence times in this case. It was done say in [6]. The result is that the decoherence time is negligibly small for macro objects.

15 Evolution of classical systems

Now, as we know that composite systems that are not entangled with anything, behave as elementary objects in potential, we can consider their evolution. Suppose, we measure a coordinate of such a system. How does this system evolve in time? For:

$$\psi(x) \sim \exp\left(\frac{-(x - \langle x \rangle)^2}{2\sigma_x^2}\right) \quad (66)$$

in the momentum space, the corresponding wave function is:

$$\phi_i(p) \sim \exp\left(-\frac{(p - \langle p \rangle)^2 \sigma_x^2}{2\hbar^2}\right) \quad (67)$$

Solving Schrödinger equation, it is easy to show that the width of the wave function is:

$$\sigma(t) = \sqrt{\sigma_0^2 + \left(\frac{\hbar t}{2m\sigma_0}\right)^2} \quad (68)$$

The characteristic time interval for the wave function to double its size therefore is:

$$t \sim \frac{2m\sigma_0^2}{\hbar} \quad (69)$$

if $\sigma_0 \ll \frac{\hbar t}{2m\sigma_0}$. t is a time interval when a classical system after measurement loses its classicity. This time is proportional to the mass of the system, and for the mass of 1 kg, and σ of 1 Å (an order for the size of an atom), it equals $10^{14}s$. During this time, the system will be measured again many times. Since the width of the wave packets almost doesn't change between the consequent measurements, we can assume that the macro-object follows about classical trajectories. For an electron however, this time interval is $10^{-15}s$, therefore its evolution cannot be described using classical trajectories alone.

For a hydrogen atom, the average time between interactions for $T \sim 100K$ is $10^{-12}s$ roughly equals the characteristic time to double its size. However, for smaller temperatures, the time between interactions is much larger leading to width in equation 68 becoming large and emergence of collective phenomena.

16 Apparatus of an intermediate size

What happens if we start scaling down our apparatus? The number of degrees of freedom N also decreases, and at some point equation 28 stops being true. In this case, the branches in equation 26 will become correlated.

Let's consider a double-slit experiment with a quantum dot near one of the slits. When the interaction of the electron with a quantum dot is negligibly weak, the electrons create a diffraction pattern on the screen. When the quantum dot

is large, two orthogonal branches are generated, and there is no interference on the screen, the pattern resembles one produced by classical objects. For the quantum dot of an intermediate size, there is some non-zero correlation between the two states of the quantum dot. The strength of the diffraction term will depend on this correlation, and be maximum when there is no quantum dot. The state vector of the system electron-Quantum Dot located near the splits is given by:

$$|e, QD\rangle_{splits} = \frac{1}{\sqrt{2}}(|e_1\rangle |QD_1\rangle + |e_2\rangle |QD_2\rangle), \quad (70)$$

where QD is the quantum dot, and indices 1 and 2 correpond to the branches. Also $\langle e_1|x\rangle = \mathcal{N}(x_1, \sigma_1)$ and $\langle e_2|x\rangle = \mathcal{N}(x_2, \sigma_2)$, where x is the axis along the plane with two slits. When traveling to the screen, the wave functions widen and get multiplied by a phase. Now, as we found the operator of evolution, we use Schrödinger equation with no potential to get the exact form of the wave functions hitting the screen. For $\sigma_1 = \sigma_2 = \sigma_0$, introducing $k_x = \frac{p_x}{\hbar}$; $k_0 = \langle k_x \rangle$; $\sigma_k = \frac{1}{\sigma_0}$, and by Fourier transforming $\mathcal{N}(x_1, \sigma_0), \mathcal{N}(x_2, \sigma_0)$, we obtain for the $\langle e_1|x\rangle$ and $\langle e_2|x\rangle$:

$$\psi(k_x, x_1^0) = \exp\left(-\frac{(k_x - k_x^0)^2}{2\sigma_k^2}\right) \exp(ik_x^0(x - x_1^0)), \quad (71)$$

$$\psi(k_x, x_2^0) = \exp\left(-\frac{(k_x - k_x^0)^2}{2\sigma_k^2}\right) \exp(ik_x^0(x - x_2^0)), \quad (72)$$

where x_1^0 and x_2^0 are the coordinates of the two slits. The phases $\exp(ik_x^0(x - x_i))$ are added to make sure that the wave functions in coordinate representation are centered around x_1^0 and x_2^0 respectively.

The probability of an electron to hit the spot with coordinate x is given by:

$$|\langle e, QD|x\rangle|_{screen}^2 = \sqrt{\frac{2\pi}{\Delta}} \left| \left(\exp\left(-\frac{(x - x_1^0)^2}{2\Delta}\right) \exp(ik_x^0(x - x_1^0)) |QD_1\rangle + \exp\left(-\frac{(x - x_2^0)^2}{2\Delta}\right) \exp(ik_x^0(x - x_2^0)) |QD_2\rangle \right) \right|^2, \quad (73)$$

where $\Delta = \sigma_0^2$. Solving Schrödinger equation, we get at time t :

$$|\langle e, QD|x\rangle|_{screen}^2 = \sqrt{\frac{2\pi}{\Delta(t)}} \left| \left(\exp\left(-\frac{(x - x_1 - \frac{\hbar k_x^0 t}{m})^2}{2\Delta(t)}\right) \exp(ik_x^0(x - x_1)) |QD_1\rangle + \exp\left(-\frac{(x - x_2 - \frac{\hbar k_x^0 t}{m})^2}{2\Delta(t)}\right) \exp(ik_x^0(x - x_2)) |QD_2\rangle \right) \right|^2, \quad (74)$$

where $\Delta(t) = \sigma_0^2 + i\frac{\hbar}{2m}t$. For simplicity, let's neglect the Gaussian terms in equation 74. This can be done if the distance between a quantum dot and the screen is much greater than the distance between the quantum dots¹⁶. After

¹⁶We need to satisfy $\frac{\hbar t}{m} \gg (x_2 - x_1)^2$. For $x_2 - x_1 \sim 1mm$, $t \gg 10^{-1}s$. Taking the average velocity of an electron to be $10^3 - 10^4 m/s$, we see that the distance between a quantum and the screen should be at least $10^3 m$. In reality, this condition is hard to achieve. But correctly taking into account these terms just complicates the final equations without adding new information

some simplifications, we get:

$$| \langle e, QD | x \rangle |_{screen}^2 = 2 + (\exp(ik_x^0(x_2 - x_1)) \langle QD_1 | QD_2 \rangle + H.C.). \quad (75)$$

From $k_x = \frac{p_x}{\hbar}$, and $p_x = p \sin \theta$, it follows that $k_x^0 = k^0 \sin \theta$, where θ is the scattering angle.

$$| \langle e, QD | x \rangle |_{screen}^2 = 2 + (\exp(ik^0(x_2 - x_1) \sin \theta) \langle QD_1 | QD_2 \rangle + H.C.). \quad (76)$$

And finally since intensities are proportional to probabilities:

$$I(x) = 2I_{1,2}(x) + I_{1,2}(\exp(ik^0(x_2 - x_1)\frac{x}{d}) \langle QD_1 | QD_2 \rangle + H.C.), \quad (77)$$

where d is the distance between the slits and the screen, and $I_{1,2} = I_1 = I_2$ is an intensity of a spot on the screen left by an electron if it goes through only one of the slits. When $\langle QD_1 | QD_2 \rangle \sim 0$, intensities add up, and when $|\langle QD_1 | QD_2 \rangle| \sim 1$, the diffraction pattern is formed:

$$I(x) = 2I_{1,2}(x) + 2I_{1,2} \cos\left(k^0(x_2 - x_1)\frac{x}{d}\right), \quad (78)$$

Generally speaking, the apparatuses can be loosely split into four categories:

1. Macroscopic apparatuses. They branch the state vectors of the composite system.
2. Intermediate size apparatuses. The branches they created are slightly correlated and the evolution of the system can be traced using Schrödinger equation
3. Observers. Observers are apparatuses that try to observe the system they are entangled with. Since there is no way to disentangle the state vector of observer and the system, we are forced to conclude that each branch has its own observer and effectively collapses the state vector of the composite system. Any observer is a macroscopic apparatus, so an observation always effectively collapses the wave function.
4. Observers of an intermediate size. The possibility to see a superposition would be possible only for an intermediate size observer who can analyze what he sees. All our senses are macroscopic so at present, such observers don't exist.

17 Conclusion

We tried to build QM starting from CM and generalizing it until the final model is consistent. We managed to justify the following axioms and naturally arrive at the following important notions.

1. Complexity of the wave function.
2. Various representations and state vectors.
3. Special place of the momentum operator in QM.

4. Space spanned by non-interacting, interacting particles and weakly interacting particles.
5. Operators corresponding to measurement apparatuses.
6. Collapse and Quantum postulate.
7. Schrödinger equation.
8. Possibility to describe composite objects with state vectors.
9. Narrowness and stability of wave function for classical objects.

The rest of the results in QM are derived from these basic set of principles in the textbooks.

Acknowledgement

I want to express my appreciation to Justin Gagnon, PhD and Vadim Asnin, PhD for comments that greatly improved the manuscript.

References

- [1] L.D. Landau & E.M. Lifshitz. "Quantum Mechanics" (volume 3 of A Course of Theoretical Physics), Pergamon Press 196
- [2] R Gomatam. "Copenhagen Interpretation of Quantum Mechanics", Philosophy of Science, 2007
- [3] E. P. Wigner. "On the quantum correction for thermodynamic equilibrium". Phys. Rev. 40 (5): 1932
- [4] Joe Gunther. "Occam's Razor", Joe Gunther Series, Vol. 10, 2007
- [5] Pathak, Anirban. "Elements of Quantum Computation and Quantum Communication", London: Taylor & Francis, 2013
- [6] J. P. Paz and W. H. Zurek, "Quantum limit of decoherence: Environment induced superselection of energy eigenstates", Phys. Rev. Lett. 82, 1999
- [7] Einstein, A; B Podolsky; N Rosen (1935-05-15). "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" Physical Review. 47, 1935
- [8] "Zur Quantenmechanik der Stoßvorgänge, Max Born, Zeitschrift für Physik", 1983.
- [9] Ehrenfest, P. "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik. Zeitschrift für Physik", 1927
- [10] Slater, J.; Verma, HC . "The Theory of Complex Spectra". Physical Review., 1293